

ANALISIS BUTIR SOAL SUMATIF BAHASA JERMAN KELAS X SEMESTER 1 SMAN 1 MOJOSARI

Oktavia Jannatul Firdaus

**Mahasiswa Program Studi S1 Pendidikan Bahasa Jerman, Fakultas Bahasa dan
Seni, Universitas Negeri Surabaya**

Oktaviaf16020094011@mhs.unesa.ac.id

Dra. Fahmi Wahyuningsih, M.Pd.

**Dosen Program Studi S1 Pendidikan Bahasa Jerman, Fakultas Bahasa dan Seni,
Universitas Negeri Surabaya**

ABSTRAK

Tes sumatif bahasa Jerman bertujuan untuk mengukur penguasaan peserta didik terhadap materi pembelajaran bahasa Jerman selama satu semester. Hal tersebut perlu dilakukan untuk mengukur kesesuaian materi dengan tujuan pembelajaran selama satu semester, sehingga perlu adanya analisis butir soal sumatif. Dalam menganalisis butir soal diperlukan validitas, reliabilitas, daya pembeda, tingkat kesukaran dan distraktor. Rumusan masalah dalam penelitian ini adalah bagaimana kelayakan soal sumatif bahasa Jerman kelas X semester 1 di SMAN 1 Mojokari dilihat dari butir soalnya. Uji validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan distraktor menggunakan Microsoft Excel. Penelitian ini menggunakan pendekatan penelitian kuantitatif deskriptif dengan jenis penelitian uji coba. Sumber data yang digunakan yaitu lembar jawaban siswa soal tes sumatif yang berjumlah 34. Sumber data yang digunakan yaitu kisi-kisi dan soal tes sumatif bahasa Jerman kelas X semester 1 yang berjumlah 40 soal pilihan ganda dengan 5 pilihan jawaban. Hasil penelitian ini menunjukkan bahwa 20 soal tersebut dinyatakan valid dan 20 soal lainnya dinyatakan tidak valid. Sehingga 50% soal tes sumatif bahasa Jerman kelas X semester 1 SMAN 1 Mojokari dinyatakan valid. Sedangkan soal tes sumatif ini dikatakan mempunyai koefisien reliabilitas rendah karena memiliki nilai 0,256. Butir soal memiliki daya pembeda sangat buruk berjumlah 6, buruk 24, cukup 9, dan baik berjumlah 1. Berdasarkan perhitungan indeks kesukaran butir soal yang tergolong mudah berjumlah 27, sedang 11, dan sukar berjumlah 2. Dari analisis distraktor banyak opsi jawaban yang dapat berfungsi dengan baik. Sehingga hasil analisis butir soal sumatif bahasa Jerman kelas X semester 1 berjumlah 18 soal yang layak dan 22 soal yang tidak layak.

Kata kunci : Analisis butir soal sumatif, validitas, reliabilitas, daya pembeda, tingkat kesukaran, distractor.

ABSTRACT

The German summative test aims to measure students' mastery of German learning material for one semester. This needs to be done to measure the suitability of the material with the learning objectives for one semester, so there is a need for summative item analysis. In analyzing the items required validity, reliability, distinguishing features, the level of difficulty and distractor. The formulation of the problem in this study is how is the feasibility of the first semester's X grade German language questions at SMAN 1 Mojokari seen from the item. Test the validity, reliability, level of difficulty, distinguishing features, and contracted using Microsoft Excel. This study uses a descriptive quantitative research approach to the type of trial research. The data source used was the student answer sheet of summative test questions totaling 34. The data source used was grating and summative test questions for German language class X semester 1 which amounted to 40 multiple choice questions with 5 answer choices. The results of this study indicate that 20 questions were declared valid and 20 other questions were declared invalid. So 50% of the X grade German summative test questions in semester 1 of SMAN 1 Mojokari are valid. While this summative test item is said to have a low reliability coefficient because it has a value of 0.256. The items have very bad distinguishing power amounted to 6, bad 24, enough 9, and good amounted to 1. Based on the calculation of success index items that are classified as easy amounted to 27, moderate 11, and difficult to number 2. From the analysis of distractors many answer options that can function well. So that the results of the analysis of the summative items in German class X semester 1 amounted to 18 feasible questions and 22 improper questions.

Keywords: Summative item analysis, validity, reliability, distinguishing features, level of difficulty, distractor.

PENDAHULUAN

Tes sumatif merupakan sebuah acuan untuk mengukur sejauh mana kemampuan peserta didik dalam memahami materi. Berdasarkan hasil diskusi, guru bahasa Jerman SMAN 1 Mojosari menyatakan bahwa tes sumatif kelas X semester 1 belum pernah dilakukan analisis butir soal. Oleh sebab itu perlu adanya penelitian mengenai analisis butir soal sumatif bahasa Jerman kelas X semester 1 di SMAN 1 Mojosari. Data penelitian yang digunakan dalam penelitian ini adalah soal sumatif bahasa Jerman yang belum diujikan kepada peserta didik kelas X semester 1 SMAN 1 Mojosari.

Rumusan masalah dalam penelitian ini adalah “bagaimana kelayakan soal sumatif bahasa Jerman kelas X semester 1 di SMAN 1 Mojosari dilihat dari butir soalnya”. Penelitian ini dibatasi bahwa Soal sumatif kelas X SMAN 1 Mojosari berjumlah 40 soal berbentuk pilihan ganda, diyakini bahwa sudah di buat kisi-kisi, penelitian ini hanya mengacu pada analisis butir soal sumatif untuk mengetahui validitas, reliabilitas, tingkat kesukaran, daya pembeda dan distraktor, untuk pedoman analisis penelitian ini berdasarkan Arikunto (2013).

METODE

Penelitian ini menggunakan pendekatan penelitian kuantitatif deskriptif dengan jenis penelitian uji coba karena dalam penelitian ini akan mengujicobakan soal sumatif bahasa Jerman kelas X semester 1 SMAN 1 Mojosari. Sedangkan data yang digunakan dalam penelitian ini berupa respon peserta didik terhadap butir soal sumatif kelas X semester 1 yang

berjumlah 40 butir soal pilihan ganda. Data diperoleh menggunakan metode tes uji coba yang dilakukan pada kelas X MIPA 1 SMAN 1 Mojosari. Data tersebut selanjutnya diolah menggunakan bantuan program Ms. Excel.

HASIL DAN PEMBAHASAN

Berdasarkan hasil analisis pada soal sumatif kelas X semester 1 terdapat 20 soal valid atau 50% dari keseluruhan soal. Butir soal sumatif memiliki reliabilitas sebesar 0,256, jika instrumen tersebut mempunyai koefisien reliabilitas rendah, sehingga dapat disimpulkan bahwa soal sumatif tersebut mempunyai koefisien reliabilitas rendah.

Butir soal yang memiliki daya pembeda sangat buruk atau negatif berjumlah 6 butir soal atau 15%, butir soal yang memiliki daya pembeda buruk berjumlah 24 butir atau 60%, butir soal yang memiliki daya pembeda cukup berjumlah 9 butir atau 22,5%, butir soal yang memiliki daya pembeda baik berjumlah 1 butir atau 2,5%, dan butir soal yang memiliki daya pembeda sangat baik tidak ada.

Berdasarkan hasil perhitungan indeks kesukaran butir soal secara manual dengan bantuan Ms. Excel, butir yang tergolong mudah berjumlah 27 butir atau 67,5%, butir soal yang tergolong sedang berjumlah 11 soal atau 27,5%,

dan butir soal yang tergolong sukar berjumlah 2 atau 5%.

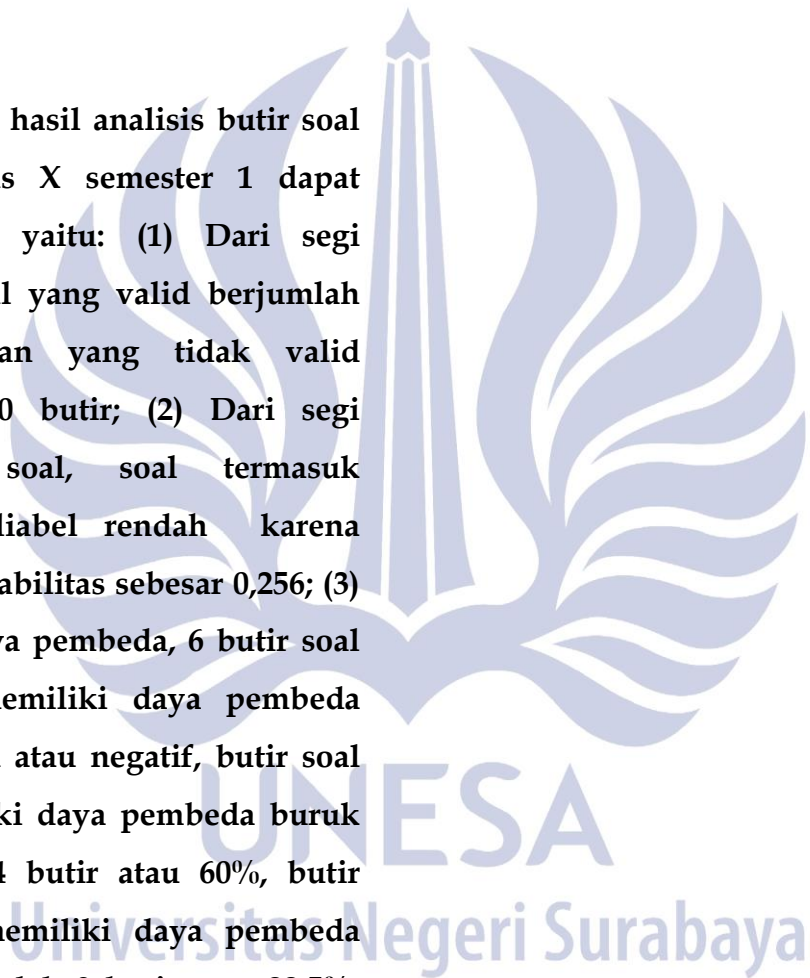
Analisis distraktor dapat dikatakan berfungsi dengan baik jika peserta didik yang menjawab lebih dari 5%,

jika yang menjawab kurang dari 5% dapat dikatakan berfungsi tidak baik. Hasil analisis distraktor dapat dikatakan soal sumatif bahasa Jerman dapat berfungsi dengan baik. Sehingga hasil analisis butir soal sumatif bahasa Jerman kelas X semester 1 berjumlah 18 soal yang layak dan 22 soal yang tidak layak.

PENUTUP SIMPULAN

Berdasarkan hasil analisis butir soal sumatif kelas X semester 1 dapat disimpulkan yaitu: (1) Dari segi validitas, soal yang valid berjumlah 30 butir dan yang tidak valid berjumlah 10 butir; (2) Dari segi reliabilitas soal, soal termasuk koefisien reliabel rendah karena memiliki reliabilitas sebesar 0,256; (3) Dari segi daya pembeda, 6 butir soal atau 15% memiliki daya pembeda sangat buruk atau negatif, butir soal yang memiliki daya pembeda buruk berjumlah 24 butir atau 60%, butir soal yang memiliki daya pembeda cukup berjumlah 9 butir atau 22,5%, butir soal yang memiliki daya pembeda baik berjumlah 1 butir atau 2,5%; (4) Dari segi tingkat kesukaran, 27 butir soal merupakan kategori mudah, 11 butir soal merupakan kategori sedang, dan 2 butir soal

merupakan kategori sukar; (5) Untuk analisis distraktor dapat dikatakan berfungsi dengan baik. Sehingga hasil analisis butir soal sumatif bahasa Jerman kelas X semester 1 berjumlah 18 soal yang layak dan 22 soal yang tidak layak.



DAFTAR PUSTAKA

- Aliaga, M. and Gunderson, B. 2002. *Interactive Statistics*. New Jersey: Prentice Hall.
- Arikunto, Suharsimi. 2013. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, Suharsimi. 2013. *Prosedur Penelitian suatu pendekatan praktik*. Jakarta: Rineka Cipta.
- Azwar, Saifuddin. 1999. *Dasar-Dasar Psikometri*. Yogyakarta: Pustaka Pelajar Offset.
- Brauns, Katrin & Schubert, Sebastian. 2008. *Qualitätssicherung von Multiple-Choice-Prüfungen*. (<https://www.elearning.uni-mainz.de/file/2014/03/Brauns-Schubert-Qualitätssicherung-von-MC-Prüfungen.pdf> diunduh pada tanggal 02 November 2019).
- Darmadi, Hamid. 2011. *Metode Penelitian Pendidikan*. Bandung: Alfabeta.
- Gamung, Yohana Magdalena. 2017. *Penerapan Teknik Skema Dalam Pembelajaran Membaca Bahasa Jerman Siswa Kelas X Ilmu Bahasa Dan Budaya SMA Negeri 3 Sidoarjo*. Skripsi. Surabaya: Universitas Negeri Surabaya.
- Husamah. 2014. *Pembelajaran Bauran (Blended Learning) Terampil Memadukan Keunggulan Pembelajaran Face-to-Face, E-Learning Office Online dan Mobile Learning*. Jakarta: Prestasi Pustaka.
- Indonesiabaik.id. 2018. *Jumlah Penggunaan Smartphone di Indonesia*. (<https://www.indonesiabaik.id/infografis/jumlah-pengguna-smartphone-di-indonesia-18>. Diakses 5 April 2019 21:11).
- Leman, Michael Andreas. 2018. *Cara Praktis Melakukan Uji Validitas Alat Ukur Penelitian*. Yogyakarta: Gosyen Publishing.
- Malawi, Ibadullah dkk. 2019. *Teori dan Aplikasi Pembelajaran Terpadu*. Magetan: CV. AE. MEDIA GRAFIKA
- Noemi, Pena Miguel dkk. 2014. *Educational Games for Learning*. Horizon Research Publishing. (<https://files.eric.ed.gov/fulltext/EJ1053979.pdf>, diakses pada 4 Oktober 2018).
- Nurgiyantoro, Burhan. 1987. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: BPFE
- Purwanto. 2009. *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.
- Purwanto, Ngalim. 1984 *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: Remadja Rosdakarya.
- Reinmann, Gabi. 2017. *Prüfungen und Forschendes Lernen*, (https://gabi-reinmann.de/wp-content/uploads/2014/12/Artikel_Pruefungen2_forschendesLernen_Dez14_Preprint.pdf diunduh pada tanggal 03 November 2019)
- Silverius, Suke. 1991. *Evaluasi Hasil Belajar dan Umpan Balik*. Jakarta: PT Grasindo.
- Sudjana, Nana. (2006). *Penilaian Hasil Belajar Proses Belajar Mengajar*. Bandung: PT.Remaja Rosdakarya.
- Sukardi. 2008. *Evaluasi Pendidikan Prinsip dan Operasionalnya*. Jakarta Timur: PT Bumi Aksara.
- Tuckman, B. W. 1975. *Measuring Educational Outcome: Fundamental of testing*. New York: Harcourt Brace Jovanovich, Inc.

ANALYSE VON SUMMATIVEN ITEMS IN DEUTSCH KLASS X SEMESTER 1 SMAN 1 MOJOSARI

Oktavia Jannatul Firdaus

Pädagogik der Deutsche Sprache, Fakultät für Sprache und Kunst, Universität
Negeri Surabaya

Oktaviaf16020094011@mhs.unesa.ac.id

Dra. Fahmi Wahyuningsih, M.Pd.

Pädagogik der Deutsche Sprache, Fakultät für Sprache und Kunst, Universität
Negeri Surabaya

AUSZUG

Der summative Deutschtest soll die Beherrschung von Deutsch als Fremdsprache für ein Semester messen. Dies muss erfolgen, um die Eignung des Materials mit den Lernzielen für ein Semester zu messen, sodass eine summative Elementanalyse erforderlich ist. Bei der Analyse der Gegenstände waren Gültigkeit, Zuverlässigkeit, Unterscheidungsmerkmale, der Schwierigkeitsgrad und Ablenkungskraft gefragt. Die Formulierung des Problems in dieser Studie ist, wie die Machbarkeit der Fragen in deutscher Sprache für das erste Semester bei SMAN 1 Mojosaari aus dem Artikel ersichtlich ist. Testen Sie die Gültigkeit, Zuverlässigkeit, den Schwierigkeitsgrad, die Unterscheidungsmerkmale und den Vertrag mit Microsoft Excel. Diese Studie verwendet einen deskriptiven quantitativen Forschungsansatz für die Art der Studienforschung. Als Datenquelle wurde das Antwortblatt der Schüler mit insgesamt 34 summativen Testfragen verwendet. Als Datenquelle wurden Gitter- und summative Testfragen für das erste Semester der Deutschklasse X verwendet, die 40 Multiple-Choice-Fragen mit 5 Antwortmöglichkeiten umfassten. Die Ergebnisse dieser Studie zeigen, dass 20 Fragen für gültig und 20 andere für ungültig erklärt wurden. Somit sind 50% der summativen Testfragen für Deutsch mit der Note X im ersten Semester von SMAN 1 Mojosaari gültig. Dieser summative Testgegenstand soll einen niedrigen Zuverlässigkeitskoeffizienten haben, da er einen Wert von 0,256 hat. Die Elemente mit sehr schlechter Unterscheidungskraft beliefen sich auf 6, schlecht auf 24, genug auf 9 und gut auf 1. Basierend auf der Berechnung von Erfolgsindexelementen, die als einfach eingestuft wurden, betrugen sie 27, mäßig 11 und schwer zu nummerieren na ja. Damit ergeben sich aus der Analyse der summativen Items im 1. Semester der Klasse X 18 realisierbare Fragen und 22 unsachgemäße Fragen.

Stichwörter: Summative Item-Analyse, Gültigkeit, Zuverlässigkeit, Unterscheidungsmerkmale, Schwierigkeitsgrad, Distraktor.

ABSTRACT

The German summative test aims to measure students' mastery of German learning material for one semester. This needs to be done to measure the suitability of the material with the learning objectives for one semester, so there is a need for summative item analysis. In analyzing the items required validity, reliability, distinguishing features, the level of difficulty and distractor. The formulation of the problem in this study is how is the feasibility of the first semester's X grade German language questions at SMAN 1 Mojokerto seen from the item. Test the validity, reliability, level of difficulty, distinguishing features, and contracted using Microsoft Excel. This study uses a descriptive quantitative research approach to the type of trial research. The data source used was the student answer sheet of summative test questions totaling 34. The data source used was grating and summative test questions for German language class X semester 1 which amounted to 40 multiple choice questions with 5 answer choices. The results of this study indicate that 20 questions were declared valid and 20 other questions were declared invalid. So 50% of the X grade German summative test questions in semester 1 of SMAN 1 Mojokerto are valid. While this summative test item is said to have a low reliability coefficient because it has a value of 0.256. The items have very bad distinguishing power amounted to 6, bad 24, enough 9, and good amounted to 1. Based on the calculation of success index items that are classified as easy amounted to 27, moderate 11, and difficult to number 2. From the analysis of distractors many answer options that can function well. So that the results of the analysis of the summative items in German class X semester 1 amounted to 18 feasible questions and 22 improper questions.

Keywords: Summative item analysis, validity, reliability, distinguishing features, level of difficulty, distractor.

HINTERGRUND

Der summative Test ist eine Referenz, um zu messen, inwieweit die Schüler in der Lage sind, das Material zu verstehen. Basierend auf den Ergebnissen der Diskussion gab der Deutschlehrer von SMAN 1 Mojosari an, dass der summative Test der Klasse X, 1. Semester, niemals mit einer Item-Analyse durchgeführt wurde. Daher besteht Forschungsbedarf zur Analyse von summativen Items im 1. Semester der Klasse X in SMAN 1 Mojosari. Bei den in dieser Studie verwendeten Forschungsdaten handelt es sich um eine summative deutsche Sprache, die für Schüler der Klasse X, 1. Semester von SMAN 1, Mojosari, nicht getestet wurde.

Die Formulierung des Problems in dieser Studie lautet: "Wie ist die Machbarkeit der Fragen in deutscher Sprache im ersten Semester der Klasse X bei SMAN 1 Mojosari von der Spitze aus gesehen?". Diese Studie ist begrenzt, dass die summativen Fragen der Klasse X SMAN 1 Mojosari 40 Multiple-Choice-Fragen belief, wird angenommen, dass das Raster gemacht wurde, bezieht sich diese Studie nur auf die Analyse der summativen Elemente zur Bestimmung der Gültigkeit, Zuverlässigkeit, Schwierigkeitsgrad, Differenzierung und Distraktor, für die analytische Anleitung dieser Studie basierend auf Arikunto (2013).

METHODE

Diese Studie verwendet einen deskriptiven quantitativen Forschungsansatz mit der Art der Studienforschung, da in dieser Studie

die summativen Fragen des ersten Semesters der Deutschklasse X von SMAN 1 Mojosari erprobt werden. Bei den in dieser Studie verwendeten Daten handelt es sich um die Antworten der Studenten auf die summativen Punkte der Klasse 1 des X. Semesters mit insgesamt 40 Multiple-Choice-Fragen. Die Daten wurden unter Verwendung einer in Klasse X MIPA 1 von SMAN 1 Mojosari durchgeführten Testmethode erhalten. Die Daten werden dann mit Hilfe von Frau Programm verarbeitet. Excel.

ERGEBNISSE UND DISKUSSION

Basierend auf den Ergebnissen der Analyse der summativen Fragen der Klasse 1 des Semesters X gibt es 20 gültige Fragen oder 50% aller Fragen. Summative Items haben eine Zuverlässigkeit von 0,256, wenn das Instrument einen Koeffizienten geringe Zuverlässigkeit hat, so dass geschlossen werden kann, dass das summative Problem einen geringen Zuverlässigkeitskoeffizienten aufweist.

Artikel mit sehr schlechter oder negativer Differenzierung sind 6 Artikel oder 15%, Artikel mit schlechter Differenzierung sind 24 Artikel oder 60%, Artikel mit ausreichender Differenzierungskraft sind 9 Artikel oder 22,5%, Artikel mit guter Differenzierung sind 1 Punkt oder 2,5%, und es gibt keine Artikel, die eine sehr gute Differenzierung aufweisen.

Basierend auf der manuellen Berechnung der Schwierigkeitsindexelemente mit Hilfe von Frau Excel sind als einfach eingestufte Elemente 27 Elemente oder

67,5%, als mittelmäßig eingestufte Elemente 11 Elemente oder 27,5% und Elemente, die als klassifiziert sind schwierig sind 2 oder 5%.

Man kann sagen, dass die Distraktoranalyse gut funktioniert, wenn Schüler mit mehr als 5% antworten. Die Ergebnisse der Distraktoranalyse können gesagt werden, dass die summative deutsche Sprache gut funktionieren kann. Damit ergeben sich aus der Analyse der summativen Items im 1. Semester der Klasse X 18 realisierbare Fragen und 22 unsachgemäße Fragen.

SCHLUSS

ZUSAMMENFASSUNG

Unterscheidung sind 24 Gegenstände oder 60%, Gegenstände mit ausreichendem Unterscheidungsvermögen sind 9 Gegenstände oder 22,5%, Gegenstände mit guter Differenzierung sind 1 Punkt oder 2,5%; (4) In Bezug auf den Schwierigkeitsgrad sind 27 Gegenstände eine einfache Kategorie, 11 Gegenstände

Basierend auf den Ergebnissen der Analyse der summativen Elemente der Klasse X in Semester 1 kann der Schluss gezogen werden, dass: (1) Gültige Fragen 30 Elemente und 10 Elemente ungültig sind; (2) In Bezug auf die Zuverlässigkeit der Fragen gehören die Fragen zu den niedrigen zuverlässigen Koeffizienten, da sie eine Zuverlässigkeit von 0,256 haben; (3) In Bezug auf das Unterscheidungsvermögen haben 6 Gegenstände oder 15% sehr schlechte oder negative Unterscheidungsmerkmale, Gegenstände mit schlechter

eine mittlere Kategorie und 2 Gegenstände eine schwierige Kategorie. (5) Für die Analyse von Distraktoren kann gesagt werden, dass sie richtig funktionieren. Damit ergeben sich aus der Analyse der summativen Items im 1. Semester der Klasse X 18 realisierbare Fragen und 22 unsachgemäße Fragen.

BIBLIOGRAPHIE

- Aliaga, M. and Gunderson, B. 2002. *Interactive Statistics*. New Jersey: Prentice Hall.
- Arikunto, Suharsimi. 2013. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, Suharsimi. 2013. *Prosedur Penelitian suatu pendekatan praktik*. Jakarta: Rineka Cipta.
- Azwar, Saifuddin. 1999. *Dasar-Dasar Psikometri*. Yogyakarta: Pustaka Pelajar Offset.
- Brauns, Katrin & Schubert, Sebastian. 2008. *Qualitätssicherung von*

Multiple-Choice-Prüfungen.

(<https://www.elearning.uni-mainz.de/file/2014/03/Brauns-Schubert-Qualitätssicherung-von-MC-Prüfungen.pdf> diunduh pada tanggal 02 November 2019).

Darmadi, Hamid. 2011. *Metode Penelitian Pendidikan*. Bandung: Alfabeta.

Gamung, Yohana Magdalena. 2017. *Penerapan Teknik Skema Dalam Pembelajaran Membaca Bahasa Jerman Siswa Kelas X Ilmu Bahasa Dan Budaya SMA Negeri 3 Sidoarjo*. Skripsi. Surabaya: Universitas Negeri Surabaya.

- Husamah. 2014. *Pembelajaran Bauran (Blended Learning) Terampil Memadukan Keunggulan Pembelajaran Face-to-Face, E-Learning Office Online dan Mobile Learning*. Jakarta: Prestasi Pustaka.
- Indonesiabaik.id. 2018. *Jumlah Penggunaan Smartphone di Indonesia*. (<https://www.indonesiabaik.id/infografis/jumlah-pengguna-smartphone-di-indonesia-18>). Diakses 5 April 2019 21:11).
- Leman, Michael Andreas. 2018. *Cara Praktis Melakukan Uji Validitas Alat Ukur Penelitian*. Yogyakarta: Gosyen Publishing.
- Malawi, Ibadullah dkk. 2019. *Teori dan Aplikasi Pembelajaran Terpadu*. Magetan: CV. AE. MEDIA GRAFIKA
- Silverius, Suke. 1991. *Evaluasi Hasil Belajar dan Umpan Balik*. Jakarta: PT Grasindo.
- Sudjana, Nana. (2006). *Penilaian Hasil Belajar Proses Belajar Mengajar*. Bandung: PT.Remaja Rosdakarya.
- Sukardi. 2008. *Evaluasi Pendidikan Prinsip dan Operasionalnya*. Jakarta Timur: PT Bumi Aksara.
- Tuckman, B. W. 1975. *Measuring Educational Outcome: Fundamental of testing*. New York: Harcourt Brace Jovanovich, Inc.
- Noemi, Pena Miguel dkk. 2014. *Educational Games for Learning*. Horizon Research Publishing. (<https://files.eric.ed.gov/fulltext/EJ1053979.pdf>, diakses pada 4 Oktober 2018).
- Nurgiyantoro, Burhan. 1987. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: BPFE
- Purwanto. 2009. *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.
- Purwanto, Ngalmim. 1984 *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: Remadja Rosdakarya.
- Reinmann, Gabi. 2017. *Prüfungen und Forschendes Lernen*, (https://gabi-reinmann.de/wp-content/uploads/2014/12/Artikel_Pruefungen2_forschendesLernen_Dez14_Preprint.pdf diunduh pada tanggal 03 November 2019)



UNESA

Universitas Negeri Surabaya